Image Coding Using Wavelet Transform

Marc Antonini, Michel Barlaud, *Member, IEEE*, Pierre Mathieu, and Ingrid Daubechies, *Member, IEEE*

Abstract—**Image compression is now essential for applications such as transmission and storage in data bases. This paper proposes a new scheme for image compression taking into account psychovisual features both in the space and frequency domains; this new method involves two steps. First, we use a *wavelet transform* in order to obtain a set of biorthogonal subclasses of images; the original image is decomposed at different scales using a pyramidal algorithm architecture. The decomposition is along the vertical and horizontal directions and maintains constant the number of pixels required to describe the image. Second, according to Shannon's rate distortion theory, the wavelet coefficients are *vector quantized* using a multiresolution codebook. Furthermore, to encode the wavelet coefficients, we propose a noise shaping bit allocation procedure which assumes that details at high resolution are less visible to the human eye. Finally, in order to allow the receiver to recognize a picture as quickly as possible at minimum cost, we present a progressive transmission scheme. It is shown that the wavelet transform is particularly well adapted to progressive transmission.**

Keywords—**Wavelet, biorthogonal wavelet, multiscale pyramidal algorithm, vector quantization, noise shaping, progressive transmission.**

I. Introduction

IN many different fields, *digitized images* are replacing conventional analog images as photograph or x-rays. The volume of data required to describe such images greatly slow transmission and makes storage prohibitively costly. The information contained in the images must, therefore, be compressed by extracting only the visible elements, which are then encoded. The quantity of data involved is thus reduced substantially.

A fundamental goal of data compression is to reduce the bit rate for transmission or storage while maintaining an acceptable fidelity or image quality. Compression can be achieved by transforming the data, projecting it on a basis of functions, and then encoding this transform. Because of the nature of the image signal and the mechanisms of human vision, the transform used must accept nonstationarity and be well localized in both the space and frequency domains. To avoid redundancy, which hinders compression, the transform must be at least biorthogonal and lastly, in order to save CPU time, the corresponding algorithm must be fast. The two-dimensional wavelet transform defined by Meyer and Lemarié [31], [24], [25], together with its implementation as described by Mallat [27], satisfies each of these conditions.

The compression method we have developed associates a wavelet transform and a vector quantization coding scheme. The wavelet coefficients are coded considering a noise shaping bit allocation procedure. This technique exploits the psychovisual as well as statistical redundancies in the image data, enabling bit rate reduction.

Section II describes the wavelet transforms used in this paper. After a quick review of wavelets in general, we explain in more detail the properties and construction of regular biorthogonal wavelet bases. We then extend this one-dimensional construction to a two-dimensional scheme with separable filters. The new coding scheme is next presented in Section III. We focus particularly in this section on the statistical properties of wavelet coefficients, on the asymptotic coding gain that can be achieved using vector quantization in the subimages, and on the optimal allocation across the subimages. Experimental results are given in Section IV for images taken within and outside of the training set.

Manuscript received February 7, 1990; revised March 26, 1991.

M. Antonini, M. Barlaud, and P. Mathieu are with LASSY 13S CNRS, Universite de Nice-Sophia Antipolis, 06560 Valbonne, France.

I. Daubechies is with AT&T Bell Laboratories, Murray Hill, NJ 07974.

IEEE Log Number 9106073.

II. Wavelets

A. A Short Review of Wavelet Analysis

Wavelets are functions generated from one single function ψ by dilations and translations

$$\psi^{a,b}(t) = |a|^{-1/2}\psi\left(\frac{t-b}{a}\right).$$

(For this introduction we assume t is a one-dimensional variable). The *mother wavelet* ψ has to satisfy $\int dx\, \psi(x) = 0$, which implies at least some oscillations. (Technically speaking, the condition on ψ should be $\int d\omega\, |\Psi(\omega)|^2\, |\omega|^{-1} < \infty$, where Ψ is the Fourier transform of ψ; if $\psi(t)$ decays faster than $|t|^{-1}$ for $t \to \infty$, then this condition is equivalent to the one above). The definition of wavelets as dilates of one function means that high frequency wavelets correspond to $a < 1$ or narrow width, while low frequency wavelets have $a > 1$ or wider width.

The basic idea of the wavelet transform is to represent any arbitrary function f as a superposition of wavelets. Any such superposition decomposes f into different scale levels, where each level is then further decomposed with a resolution adapted to the level. One way to achieve such a decomposition writes f as an integral over a and b of $\psi^{a,b}$ with appropriate weighting coefficients [22]. In practice, one prefers to write f as a discrete superposition (sum rather than integral). Therefore, one introduces a discre-

tization, $a = a_0^m$, $b = nb_0a_0^m$, with $m, n \in \mathbb{Z}$, and $a_0 > 1$, $b_0 > 0$ fixed. The wavelet decomposition is then

$$f = \sum c_{m,n}(f)\psi_{m,n} \tag{1}$$

with $\psi_{m,n}(t) = \psi^{a_0^m, nb_0a_0^m}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0)$. Decompositions of this type were studied in [14], [15]. For $a_0 = 2$, $b_0 = 1$ there exist very special choices of ψ such that the $\psi_{m,n}$ constitute an orthonormal basis, so that

$$c_{m,n}(f) = \langle \psi_{m,n}, f \rangle = \int dx\, \psi_{m,n}(x) f(x)$$

in this case. Different bases of this nature were constructed by Stromberg [36], Meyer [31], Lemarié [24], Battle [7], and Daubechies [16]. All these examples correspond to a multiresolution analysis, a mathematical tool invented by Mallat [27], which is particularly well adapted to the use of wavelet bases in image analysis, and which gives rise to a fast computation algorithm.

In a multiresolution analysis, one really has *two* functions: the mother wavelet ψ and a *scaling function* ϕ. One also introduces dilated and translated versions of the scaling function, $\phi_{m,n}(x) = 2^{-m/2}\phi(2^{-m}x - n)$. For fixed m, the $\phi_{m,n}$ are orthonormal. We denote by V_m the space spanned by the $\phi_{m,n}$; these spaces V_m describe successive approximation spaces, $\cdots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots$, each with resolution 2^m. For each m, the $\psi_{m,n}$ span a space W_m which is exactly the orthogonal complement in V_{m-1} of V_m; the coefficients $\langle \psi_{m,n}, f \rangle$, therefore, describe the information lost when going from an approximation of f with resolution 2^{m-1} to the coarser approximation with resolution 2^m. All this is translated into the following algorithm for the computation of the $c_{m,n}(f) = \langle \psi_{m,n}, f \rangle$ (for more details, see [27]):

$$\begin{aligned} c_{m,n}(f) &= \sum_k g_{2n-k}a_{m-1,k}(f) \\ a_{m,n}(f) &= \sum_k h_{2n-k}a_{m-1,k}(f) \end{aligned} \tag{2}$$

where $g_l = (-1)^l h_{-l+1}$ and $h_n = 2^{1/2}\int dx\, \phi(x-n)\phi(2x)$. In fact the $a_{m,n}(f)$ are coefficients characterizing the projection of f onto V_m. If the function f is given in sampled form, then one can take these samples for the highest order resolution approximation coefficients $a_{0,n}$, and (2) describes a subband coding algorithm on these sampled values, with low-pass filter h and high-pass filter g. Because of their association with orthonormal wavelet bases, these filters give exact reconstruction, i.e.:

$$a_{m-1,l}(f) = \sum_n [h_{2n-l}a_{m,n}(f) + g_{2n-l}c_{m,n}(f)]. \tag{3}$$

Most of the orthonormal wavelet bases have infinitely supported ψ, corresponding to filters h and g with infinitely many taps. The construction in [16] gives ψ with finite support, and therefore, corresponds to FIR filters. It follows that the orthonormal bases in [16] correspond to a subband coding scheme with exact reconstruction property, using the same FIR filters for reconstruction as for decomposition. Such filters are well known since the work of Smith and Barnwell [35] and of Vetterli [37]. The extra ingredient in the orthonormal wavelet decomposition is that it writes the signal to be decomposed as a superposition of reasonably *smooth* elementary building blocks. The filters must satisfy the additional condition:

$$\prod_{k=1}^{\infty} H(2^{-k}\xi)$$

decay faster than $C(1 + |\xi|)^{-\epsilon - 0.5}$ as $|\xi| \to \infty$, for some $\epsilon > 0$, where

$$H(\xi) = 2^{-1/2}\sum_n h_n e^{-jn\xi}.$$

This extra regularity requirement is usually not satisfied by the exact reconstruction filters in the ASSP literature.

B. Applications of Wavelet Bases to Image Analysis

1) Biorthogonal Wavelet Bases: Since images are mostly smooth (except for occasional edges) it seems appropriate that an exact reconstruction subband coding scheme for image analysis should correspond to an orthonormal basis with a reasonably smooth mother wavelet. In order to have fast computation, the filters should be short (short filters lead to less smoothness, however, so they cannot be too short). On the other hand it is desirable that the FIR filters used be linear phase, since such filters can be easily cascaded in pyramidal filter structures without the need for phase compensation. Unfortunately, there are no nontrivial orthonormal linear phase FIR filters with the exact reconstruction property [35], regardless of any regularity considerations. The only symmetric exact reconstruction filters are those corresponding to the Haar basis, i.e., $h_0 = h_1 = 2^{1/2}$ and $g_0 = -g_1 = 2^{1/2}$, with all other h_n, $g_n = 0$.

One can preserve linear phase (corresponding to symmetry for the wavelet) by relaxing the orthonormality requirement, and using biorthogonal bases. It is then still possible to construct examples where the mother wavelets have arbitrarily high regularity.

In such a scheme, we still decompose as in (2), but reconstruction becomes

$$a_{m-1,l}(f) = \sum_n [\tilde{h}_{2n-l}a_{m,n}(f) + \tilde{g}_{2n-l}c_{m,n}(f)] \tag{4}$$

where the filters $\tilde{h}$, $\tilde{g}$ may be different from h, g. In order to have exact reconstruction, we impose:

$$\begin{aligned} \tilde{g}_n &= (-1)^n h_{-n+1} \\ g_n &= (-1)^n \tilde{h}_{-n+1} \end{aligned} \qquad \sum_n h_n \tilde{h}_{n+2k} = \delta_{k,0}. \tag{5}$$

So far, we have not performed anything differently from the usual exact reconstruction subband coding schemes with synthesis filters different from the decomposition filters. If the filters satisfy the additional condition that:

$$\prod_{k=1}^{\infty} \tilde{H}(2^{-k}\xi) \quad \text{and} \quad \prod_{k=1}^{\infty} H(2^{-k}\xi) \tag{6a}$$

decay faster than $C(1 + |\xi|)^{-\epsilon-0.5}$ as $|\xi| \to \infty$, for some $\epsilon > 0$, where

$$\tilde{H}(\xi) = 2^{-1/2} \sum_n \tilde{h}_n e^{-jn\xi} \quad H(\xi) = 2^{-1/2} \sum_n h_n e^{-jn\xi} \tag{6b}$$

then we can give the following interpretation to (2) and (4). Define functions ϕ and $\tilde{\phi}$ by

$$\phi(x) = \sum_n h_n \phi(2x - n) \quad \text{and} \quad \tilde{\phi}(x) = \sum_n \tilde{h}_n \tilde{\phi}(2x - n).$$

Their Fourier transforms are exactly the infinite products (6a), and they are, therefore, well-defined square integrable functions, compactly supported if the filters h and $\tilde{h}$ are FIR. Define also

$$\psi(x) = \sum_n g_n \phi(2x - n) \quad \text{and} \quad \tilde{\psi}(x) = \sum_n \tilde{g}_n \tilde{\phi}(2x - n).$$

Then, the $a_{m,n}(f)$ and $c_{m,n}(f)$ in (2) can be rewritten as:

$$a_{m,n}(f) = \langle \phi_{m,n}, f \rangle = 2^{-m/2} \int dx\, \phi_{m,n}(x) f(x)$$

$$c_{m,n}(f) = \langle \psi_{m,n}, f \rangle = 2^{-m/2} \int dx\, \psi_{m,n}(x) f(x)$$

and reconstruction is simply:

$$f = \sum_{m,n} \langle \psi_{m,n}, f \rangle \tilde{\psi}_{m,n}. \tag{7}$$

The filter bank structure with the associating wavelets and scaling functions is depicted on the following subband coding scheme (Fig. 1).

If the infinite products in (6a) decay even faster than imposed above, then ϕ and $\tilde{\phi}$ and consequently ψ and $\tilde{\psi}$ will be reasonably smooth. Note that (7) is very similar to the orthonormal decomposition described in Section II-A; the only difference is that the expansion of f with respect to the basis $\tilde{\psi}_{m,n}$ uses coefficients computed via the dual basis $\psi_{m,n}$ with $\tilde{\psi}$ different from ψ. This interpretation is not possible for all exact reconstruction subband coding schemes; in particular, convergence of the infinite products (6a) is only possible if

$$\sum_n h_n = 2^{1/2} \quad \text{and} \quad \sum_n \tilde{h}_n = 2^{1/2}.$$

Moreover, (7) can only hold if

$$\sum_n (-1)^n h_n = 0 \quad \text{and} \quad \sum_n (-1)^n \tilde{h}_n = 0.$$

Most exact reconstruction subband coding schemes do not satisfy these conditions.

Biorthogonal bases of wavelets have recently been constructed, with regularity simultaneously but independently, by Cohen, Daubechies and Feauveau [12] and by Herley and Vetterli [38]. Reference [12] contains a detailed mathematical study, with proofs that, under the conditions stated above, the wavelets do indeed constitute numerically stable bases (Riesz bases) and a discussion of necessary and sufficient conditions for regularity. In [18]

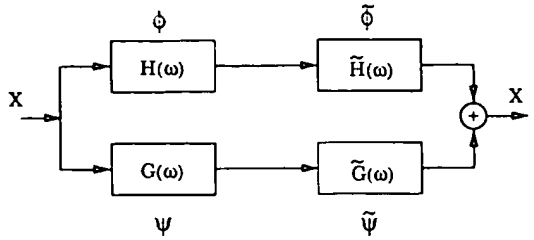


Fig. 1. Filter bank structure and the associating wavelets.

Feauveau explores the construction from the point of view of multiresolution spaces rather than from the filters. Basically one has two hierarchies of spaces in the biorthogonal case, each corresponding to one pair of filters.

It is shown in [12] that arbitrarily high regularity can be achieved by both ψ and $\tilde{\psi}$, provided one chooses sufficiently long filters. In particular, if the functions ψ and $\tilde{\psi}$ are, respectively, $(k - 1)$ and $(\tilde{k} - 1)$ times continuously differentiable, then the trigonometric polynomials $H(\xi)$ and $\tilde{H}(\xi)$ have to be divisible by $(1 + e^{-j\xi})^k$ and $(1 + e^{-j\xi})^{\tilde{k}}$, respectively, so that the length of the corresponding filters h, $\tilde{h}$ has to exceed k, $\tilde{k}$.

By (5), divisibility of $\tilde{H}(\xi)$ by $(1 + e^{-j\xi})^{\tilde{k}}$ means that ψ will have $\tilde{k}$ consecutive moments zero:

$$\int dx\, x^l \psi(x) = 0, \qquad \text{for} \quad l = 0, 1, \cdots, \tilde{k} - 1.$$

For more details concerning this discussion, see [12].

It is well known (and it can easily be checked by using Taylor expansions) that if ψ has $\tilde{k}$ moments zero, then the coefficients $\langle \psi_{m,n}, f \rangle$ will represent functions f, which are $\tilde{k}$ times differentiable, with a high compression potential (many coefficients will be negligibly small).

Many examples of biorthogonal wavelet bases with reasonably regular ψ and $\tilde{\psi}$ can be constructed; for our applications, the regularity of the elementary building blocks $\tilde{\psi}_{m,n}$, which is linked to the number of zero moments of ψ, is more important than the regularity of the $\psi_{m,n}$ or the number of zero moments of $\tilde{\psi}$. Within the limits imposed by the support widths, we will, therefore, try to choose $\tilde{k}$ as large as possible.

In terms of trigonometric polynomials $H(\xi)$ and $\tilde{H}(\xi)$, the exact reconstruction requirement condition on h and $\tilde{h}$ given in (5) reduces to (for symmetric filters)

$$H(\xi)\tilde{H}(\xi) + H(\xi + \pi)\tilde{H}(\xi + \pi) = 1. \tag{8}$$

Together with divisibility of H and $\tilde{H}$, respectively, by $(1 + e^{-j\xi})^k$ and $(1 + e^{-j\xi})^{\tilde{k}}$, this leads to (see [12])

$$H(\xi)\tilde{H}(\xi) = \cos(\xi/2)^{2l} \left[\sum_{p=0}^{l-1} \binom{l - 1 + p}{p} \cdot \sin(\xi/2)^{2p} + \sin(\xi/2)^{2l} R(\xi) \right] \tag{9}$$

where $R(\xi)$ is an odd polynomial in $\cos(\xi)$, and where $2l = k + \tilde{k}$ (symmetry of h and $\tilde{h}$ forces $k + \tilde{k}$ to be even).

TABLE I
FILTER COEFFICIENTS FOR THE SPLINE FILTERS WITH $l = 3$, $k = 4$, $\tilde{k} = 2$

n	0	±1	±2	±3	±4
$2^{-1/2}h_n$	45/64	19/64	−1/8	−3/64	3/128
$2^{-1/2}\tilde{h}_n$	1/2	1/4	0	0	0

Many examples are possible. We have studied in particular the following three examples, which belong to three different families.

2) Spline Filters: One can choose, e.g., $R \equiv 0$, with $\tilde{H}(\xi) = \cos(\xi/2)^{\tilde{k}} e^{-j\kappa\xi/2}$ where $\kappa = 0$ if $\tilde{k}$ is even, $\kappa = 1$ if $\tilde{k}$ is odd. This corresponds to the filters called "spline filters" in [12] (because the corresponding function $\tilde{\phi}$ is a B-spline function) or "binomial filters" in [38] (because the $\tilde{h}$ are simply binomial coefficients). It then follows that:

$$H(\xi) = \cos(\xi/2)^{2l-\tilde{k}} e^{j\kappa\xi/2} \cdot \left[\sum_{p=0}^{l-1} \binom{l-1+p}{p} \sin(\xi/2)^{2p}\right]. \quad (10)$$

We have looked at one example from this family; it corresponds to $l = 3$, $\tilde{k} = 2$. The coefficients h_n and $\tilde{h}_n$ are listed in Table I; the corresponding scaling functions and wavelets are plotted in Fig. 2.

It is clear that the two filters in the first example have very uneven length. This is typical for all the examples in this family of "spline filters."

3) A Spline Variant with Less Dissimilar Lengths: This family still uses $R \equiv 0$ in (9), but factorizes the right-hand side of (9), breaking up the polynomial of degree $l-1$ in $\sin(\xi/2)$ into a product of two polynomials in $\sin(\xi/2)$ with real coefficients, one to be allocated to H, the other to $\tilde{H}$, so as to make the lengths of h and $\tilde{h}$ as close as possible.

The example presented here is the "smallest" one in this family (shortest h and $\tilde{h}$); it corresponds to $l = 4$ and $k = 4$. The filter coefficients are listed in Table II; the corresponding scaling functions and wavelets are plotted in Fig. 3.

Note that, unlike examples 1 and 3 where the $2^{-1/2}h_n$, $2^{-1/2}\tilde{h}_n$ are rational, the entries in Table II are truncated decimal expansions of irrational numbers. The functions ϕ in examples 1 and 2 look very similar (compare Figs. 2(a) and 3(a)); a more detailed analysis shows that the one in example 2 is more regular, however. Both correspond to 4 vanishing moments for $\tilde{\psi}$.

4) Filters Close to Orthonormal Filters: Finally, there exist many examples for which $R \neq 0$. In particular there exists a special choice of R for which the two filters are very close to each other, and both very close to an orthonormal wavelet filter.

Surprisingly, for the first example of this series, one of the two filters is a Laplacian pyramid filter proposed in [9]. It corresponds to $l = 2$, $k = 2$ and $R(\xi) = 48\cos(\xi)/175$. The filter coefficients are listed in Table III; the corresponding scaling functions and wavelets are plotted in Fig. 4. It is clear that the scaling functions ϕ and $\tilde{\phi}$ are very similar, corresponding to very similar ψ and $\tilde{\psi}$. Note that in this case, the filter coefficients are again rational.

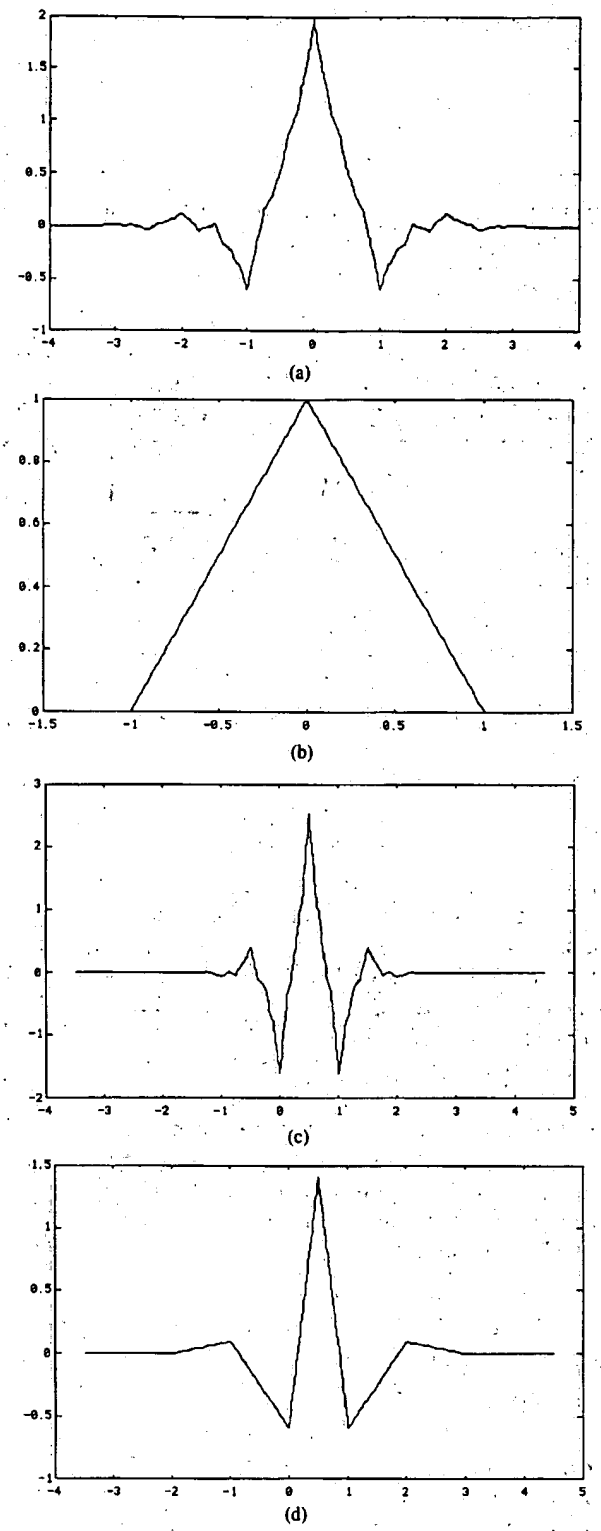


Fig. 2. Scaling functions ϕ, $\tilde{\phi}$ and wavelets ψ, $\tilde{\psi}$ for example 1 (spline filters with $l = 3$; $k = 4$, $\tilde{k} = 2$). (a) Scaling function ϕ. (b) Scaling function $\tilde{\phi}$. (c) Wavelet ψ. (d) Wavelet $\tilde{\psi}$.

TABLE II
FILTER COEFFICIENTS FOR THE SPLINE VARIANT WITH LESS DISSIMILAR LENGTHS, WITH $l = 4 = k$, $\tilde{k} = 4$

n	0	±1	±2	±3	±4
$2^{-1/2}h_n$	0.602 949	0.266 864	−0.078 223	−0.016 864	0.026 749
$2^{-1/2}\tilde{h}_n$	0.557 543	0.295 636	−0.028 772	−0.045 636	0

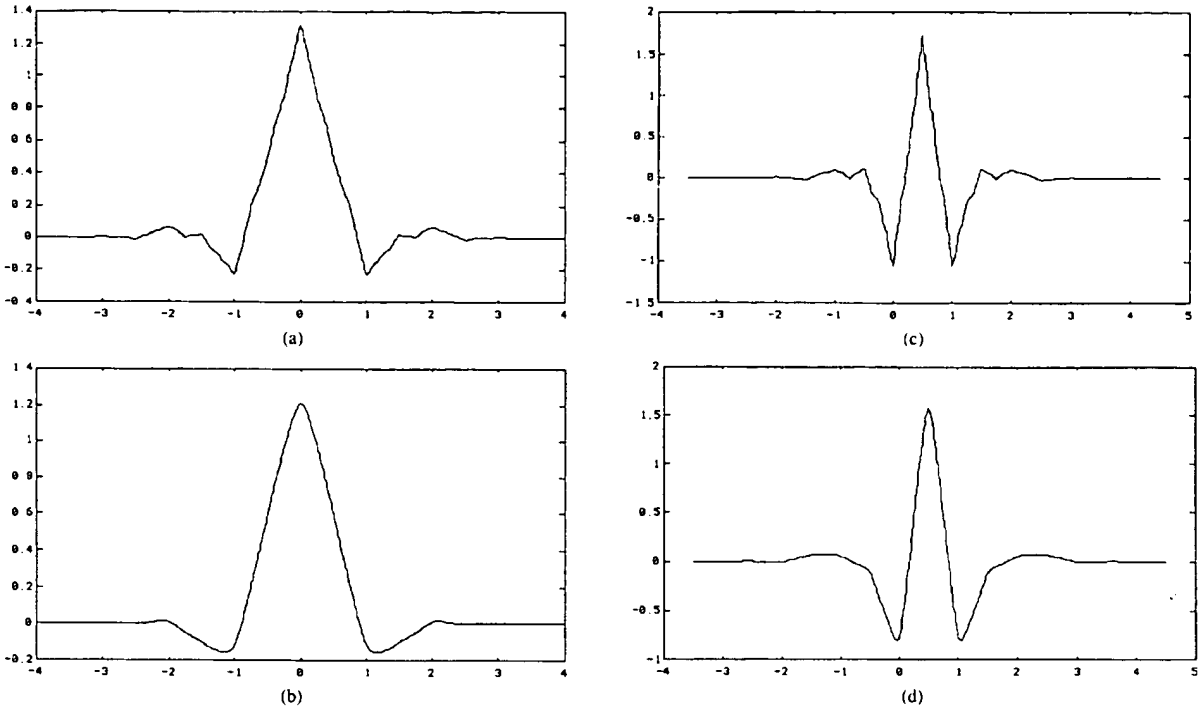


Fig. 3. Scaling functions ϕ, $\tilde{\phi}$ and wavelets ψ, $\tilde{\psi}$ for example 2 (spline variant with less dissimilar lengths; $l = 4 = k$, $\tilde{k} = 4$). (a) Scaling function ϕ. (b) Scaling function $\tilde{\phi}$. (c) Wavelet ψ. (d) Wavelet $\tilde{\psi}$.

TABLE III
FILTER COEFFICIENTS FOR EXAMPLE 3. THE ENTRIES ARE RATIONAL, AND THE TWO FILTERS ARE VERY CLOSE. THE h-FILTER COINCIDES WITH A LAPLACIAN PYRAMID FILTER PROPOSED IN [9]. IN THIS CASE $l = 2 = k$, $\tilde{k} = 2$

n	0	±1	±2	±3	±4
$2^{-1/2}h_n$	0.6	0.25	−0.05	0	0
$2^{-1/2}\tilde{h}_n$	17/28	73/280	−3/56	−3/280	0

The two biorthogonal filters in this example are both close to an orthonormal wavelet filter of length 6 constructed in [17], where it was called a "coiflet." Being an orthonormal wavelet filter, the coiflet is nonsymmetric. The filters in this example are shorter than in examples 1 and 2, but k is also smaller. The next example in this family corresponds to $k = 4$ (and $l = 4$); the filters h and $\tilde{h}$ then have length 9 and 15; they are both close to a coiflet of length 12.

5) Extension to the Two-Dimensional Case: There exist various extensions of the one-dimensional wavelet transform to higher dimensions. We follow Mallat [27] and use a two-dimensional wavelet transform in which horizontal and vertical orientations are considered preferential.

In two-dimensional wavelet analysis one introduces, like in the one-dimensional case, a scaling function $\phi(x, y)$ such that:

$$\phi(x, y) = \phi(x)\phi(y) \tag{11}$$

where $\phi(x)$ is a one-dimensional scaling function.

Let $\psi(x)$ be the one-dimensional wavelet associated with the scaling function $\phi(x)$. Then, the three two-dimensional wavelets are defined as:

$$\begin{aligned} \psi^H(x, y) &= \phi(x)\psi(y) \\ \psi^V(x, y) &= \psi(x)\phi(y) \\ \psi^D(x, y) &= \psi(x)\psi(y). \end{aligned} \tag{12}$$

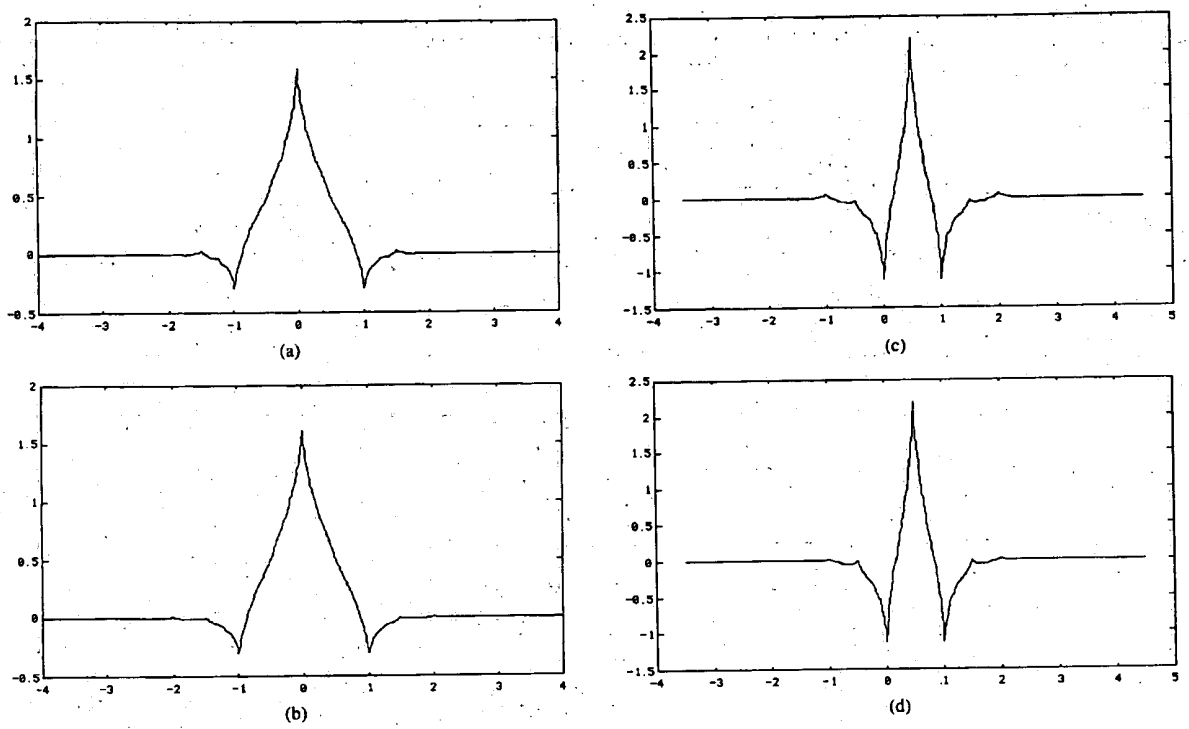


Fig. 4. Scaling functions ϕ, $\tilde{\phi}$ and wavelets ψ, $\tilde{\psi}$ for example 3 (biorthogonal filters close to an orthonormal wavelet filter, $l = 2 = k$, $\tilde{k} = 2$). (a) Scaling function ϕ. (b) Scaling function $\tilde{\phi}$. (c) Wavelet ψ. (d) Wavelet $\tilde{\psi}$.

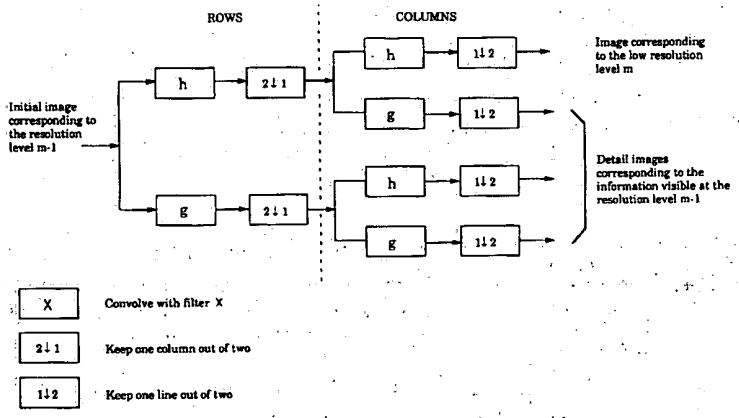


Fig. 5. One stage in a multiscale image decomposition.

Fig. 5 represents one stage in a *multiscale pyramidal* decomposition of an image: wavelet coefficients of the image are computed, as in the one-dimensional case (Sections II-A and II-B.1), using a subband coding algorithm. The filters h and g are one-dimensional filters. This decomposition provides subimages corresponding to different resolution levels and orientations (see Fig. 6). The reconstruction scheme of the image is presented Fig. 7.

To compare the three different filters presented in this paper, we have decomposed the image Lena (Fig. 16) with each of these filters. The results are presented in Fig. 8. In Fig. 8(a) we can see the normalized detail subimages at different resolution levels $m = 1$, $m = 2$, and $m = 3$ (wavelet coefficients) and in Fig. 8(b) the low resolution level subimages.

III. Image Coding Application

A. Statistical Properties of Wavelet Coefficients

The performance of a coder used for a given resolution and direction can be determined by the statistics of the corresponding subimage, i.e., its probability density function (PDF).

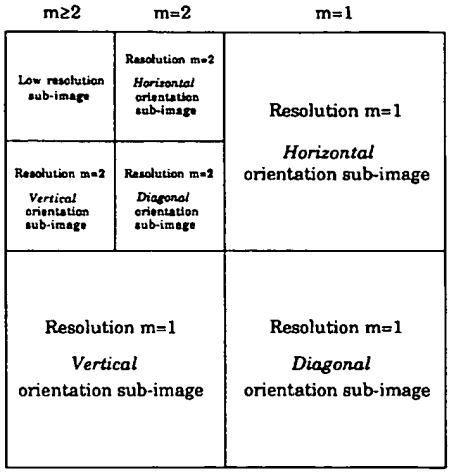


Fig. 6. Image decomposition.

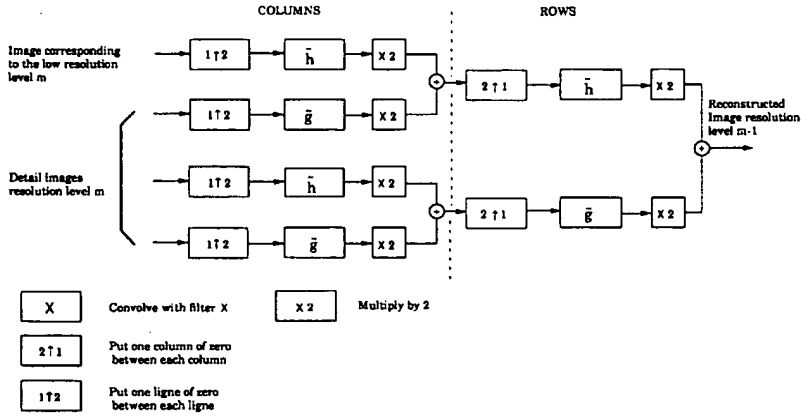


Fig. 7. One stage in a multiscale image reconstruction.

A typical PDF and different approximations are given in Fig. 9, where we plot the true PDF for resolution level $m = 1$ and direction d = vertical together with three model functions: a Gaussian, a Laplacian, and an intermediate function, the so-called generalized Gaussian [2].

This generalized Gaussian law is given explicitly by

$$p_{m,d}(x) = a_{m,d} \exp\left(-|b_{m,d}x|^{r_{m,d}}\right)$$

with

$$a_{md} = \frac{b_{m,d} r_{m,d}}{2\Gamma\left(\frac{1}{r_{m,d}}\right)} \quad \text{and} \quad b_{m,d} = \frac{1}{\sigma_{m,d}} \frac{\Gamma\left(\frac{3}{r_{m,d}}\right)^{1/2}}{\Gamma\left(\frac{1}{r_{m,d}}\right)^{1/2}} \tag{13}$$

where $\sigma_{m,d}$ is the standard deviation of the subimage (m, d), and $\Gamma(\cdot)$ is the usual Gamma function.

The general formula (13) contains the other two examples as particular cases:

- $r_{m,d} = 2$ leads to the well-known Gaussian PDF;
- $r_{m,d} = 1$ leads to a Laplacian PDF.

The variance of this approximation model is set equal to the variance of the corresponding subimage. Thus the parameter $r_{m,d}$ is computed in order to match the real PDF using the well-known chi-squared test. In this case the optimum parameter was 0.7. Other experiments for other resolutions (except the lowest resolution) lead to very similar results.

We can see in Fig. 9 that the real PDF (scale $m = 1$ and vertical orientation) is closely approximated by a generalized Gaussian law with parameter $r_{1,v} = 0.7$.

B. Encoding of Wavelet Coefficients Using Vector Quantization

Different techniques involving vector or scalar quantization can be used to encode wavelet coefficients.

According to Shannon's rate distortion theory, better results are always obtained when vectors rather than sca-

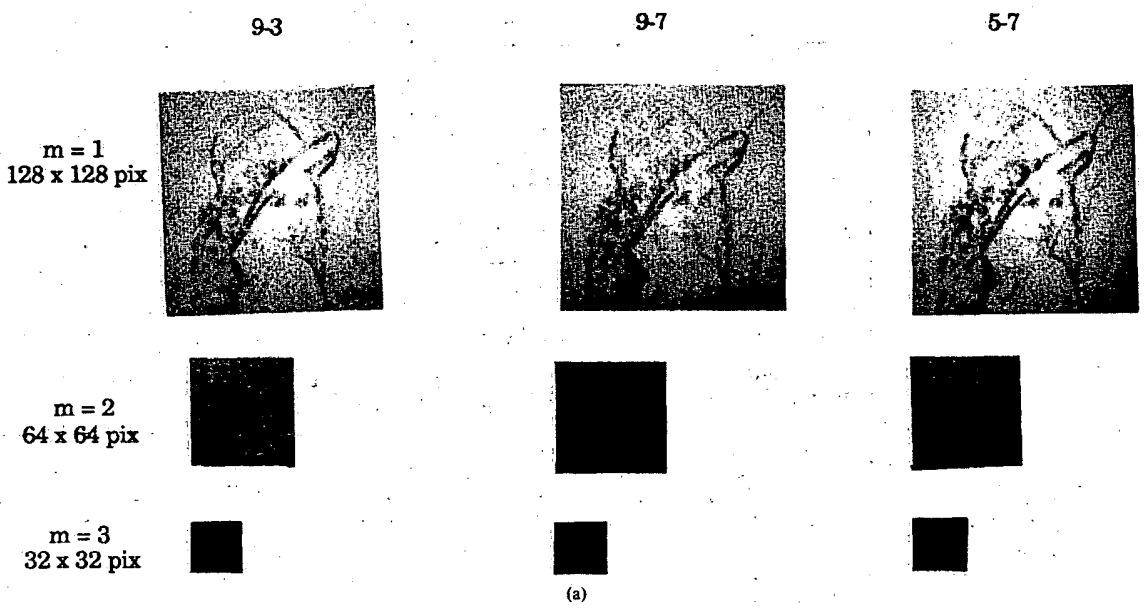


Fig. 8. Comparison among the different subimages. (a) Comparison among the normalized detail subimages. (b) Comparison among the low resolution level subimages.

lars are encoded. Therefore, the present application uses vector quantization.

1. Principle of Vector Quantization: Developed recently by Gersho and Gray (1980) [20], [21], vector quantization has proven to be a powerful tool for digital image compression [4], [29], [30], [32], [39]. The principle involves encoding a sequence of samples (vector) rather than encoding each sample individually. Encoding is performed by approximating the sequence to be coded by a vector belonging to a catalogue of shapes, usually known as a codebook.

The codebook is created and optimized using the well-known Linde-Buzo-Gray (LBG) [26] classification algorithm with a mean squared error (MSE) criterion. This algorithm is designed to perform a classification based on a training set comprised of vectors belonging to different images; it converges iteratively toward a locally optimal codebook.

Each of the vectors in the codebook is indexed. At the encoding stage, the index of the vector in the codebook most closely describing (in terms of MSE criterion) the sample set to be encoded is selected to represent this set. Of course, in order to reconstruct the sample set, the decoder must have the same codebook as the coder.

The encoding/decoding scheme depicted in Fig. 10 was proposed in [29] and [30] for orthonormal wavelets.

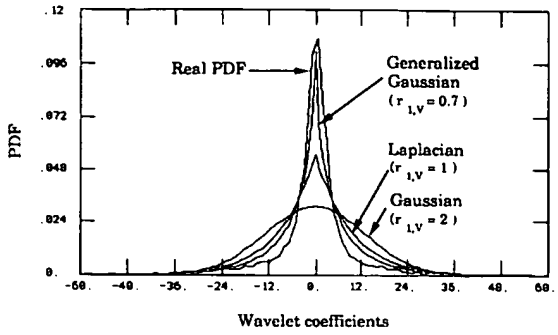


Fig. 9. Real PDF of subimage at scale $m = 1$ for vertical orientation, and its different approximations.

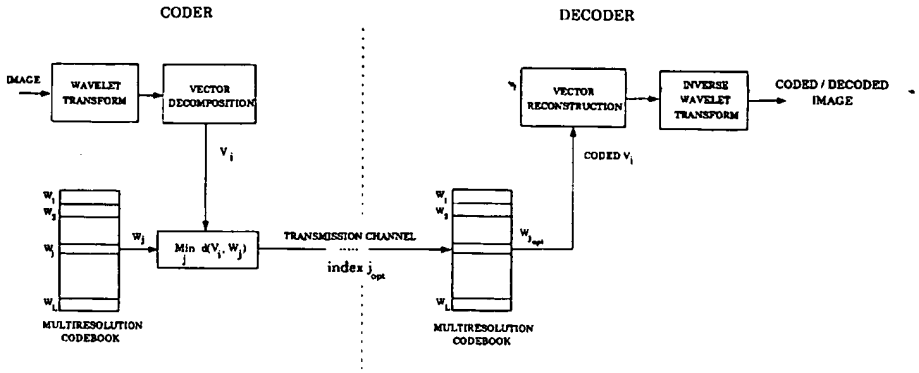


Fig. 10. Encoding/decoding scheme.

2) Comparative Performances of Vector Quantization (VQ) and Scalar Quantization (SQ): According to [3], [13], [19], [43], [30] the asymptotic lower bound distortion gain obtained when VQ, rather than SQ, is applied to a subimage is expressed as:

$$G_{m,d}^{VQ} \geq \frac{2^{-c}}{(c+1)A(k_{m,d}, c)} \times \frac{\left[\int [p_{m,d}(x)]^{1/(c+1)}\, dx\right]^{(c+1)}}{\left[\int [p_{m,d}(x)]^{k_{m,d}/(c+k_{m,d})}\, dx\right]^{(c+k_{m,d})}} \tag{14}$$

for a subimage corresponding to resolution m and direction d. $p_{m,d}(x)$ is the PDF of wavelet coefficients of the subimage with resolution m and direction d.

Here, the MSE criterion is used as a distortion measure ($c = 2$). The values of $A(k_{m,d}, 2)$ used are the upper bounds of the MSE computed and tabulated by Conway and Sloane for vector size $k_{m,d}$ [13]. This formula gives an indication of the minimum theoretical gain that can be obtained.

However, this approximation is valid only for small quantization errors, i.e., for a high bit rate $R_{m,d}$. Thus the gain $G_{m,d}^{VQ}$ only gives here an asymptotic indication.

In Fig. 11, the curves of $G_{m,d}^{VQ}$ are plotted as a function of the vector dimension $k_{m,d}$ for the Laplacian, Gaussian,

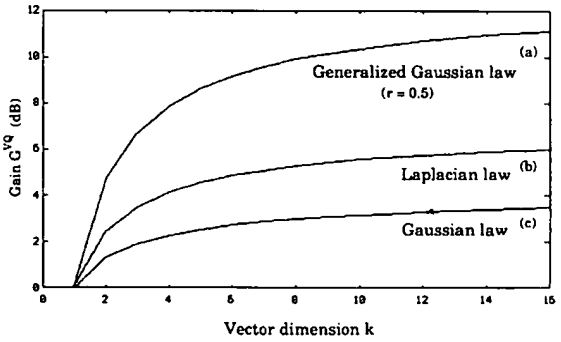


Fig. 11. Asymptotic lower bound distortion gain $G_{m,d}^{VQ}$ = function $(k_{m,d})$.

and generalized Gaussian approximation laws, and for a subimage at scale $m = 1$ and vertical orientation. Experimental results are closely matched by the theoretical results for a generalized Gaussian law with $r_{m,d} = 0.7$ except for the lower subband. Therefore, all computations based on this approximation law show that, in each subband, VQ outperforms SQ (see Fig. 11).

In summary VQ performs better for coding wavelet coefficients.

3) Generation of a Multiresolution Codebook: The preceding paragraph explained why VQ outperforms other methods. Nonetheless, major problems are encountered in the VQ of images.

- It is impossible to create a universal codebook (efficient for each image to be encoded).

- The LBG algorithm smooths high frequencies (loss of resolution).
- There is a trade-off between low distortion and high compression rate (computational cost).
- It is not easy to take into account the properties of the human visual system [28], [33].

The use of the wavelet transform (i.e., multiresolution) is one way of overcoming these different problems.

The wavelet decomposition of an image enables the generation of a codebook containing two-dimensional vectors for *each resolution level and preferential direction* (horizontal, vertical, and diagonal). Each of these subcodebooks (see Fig. 12) is generated using the LBG algorithm.

- The training set is comprised of vectors belonging to different images corresponding to the resolution and orientation under consideration.
- The initial codebook is generated by splitting the centroid (center of gravity) of this training set [21].

A multiresolution codebook can thus be obtained by assembling all of these resulting subcodebooks. Each subcodebook has a low distortion level and contains few words, which clearly facilitates the search for the best coding vector; the coding computational load is reduced, because only the appropriate subcodebook (resolution direction) of the multiresolution codebook is checked for each input vector. In addition, the quality of the coded image is better. The multiresolution codebook is depicted in Fig. 12.

Global codebook design has drawbacks in that it results in edge smoothing while the proposed method preserves edges. In fact, each subcodebook contains the shape of the wavelet coefficients which are most highly representative in terms of the MSE criterion.

Since the spatial and frequency aspects of the image are taken into account in the wavelet decomposition, the classification and search during the encoding of a subimage vector can be achieved using a simple criterion such as least mean squares. This frees us from using distortion measurements such as weighted least mean squares or other measurements involving perceptual factors. These algorithms are indeed costly in computation time.

C. Optimal Bit Allocation

Multiresolution exploits the eye's masking effects, and therefore, enables us to refine and select the type of coding according to the resolution level and the contour orientation. Although a flat noise shape minimizes the MSE criterion, it is generally not optimal for a subjective quality of image. To apply *noise shaping* across the VQ subimages, we define a total weighted MSE distortion $D_T^*(R_T)$ ((17)) for a total bit rate R_T ((18)).

Let us define $D_{m,d}(R_{m,d})$ the average distortion in the coding of the subimage (m, d) for $R_{m,d}$ bits per pixel:

$$D_{m,d}(R_{m,d}) = E(|x - q(x)|^c) = d(x, q(x)) \qquad c \geq 1 \tag{15}$$

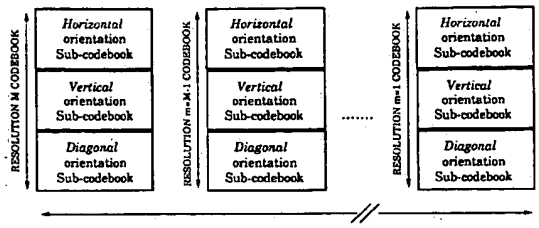


Fig. 12. Multiresolution codebook.

for all coefficients x belonging to the subimage, $q(x)$ being the quantization of x.

Total distortion of the image for a total rate of R_T bits per pixel is then given by:

$$D_T(R_T) = \frac{1}{2^{2M}} D_M^{SQ}(R_M^{SQ}) + \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} D_{m,d}(R_{m,d}) \tag{16}$$

where $D_M^{SQ}(R_M^{SQ})$ corresponds to the distortion in the subimage of lowest resolution M (texture subimage).

The problem of finding an optimal bit assignment (in bits per pixel) for each subimage vector quantizer is then formulated as:

$$\min_{R_{m,d}} \left[D_T^*(R_T) = \frac{1}{2^{2M}} D_M^{SQ}(R_M^{SQ}) + \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} D_{m,d}(R_{m,d}) \times B_{m,d} \right] \tag{17}$$

$$\textbf{subject to: } R_T = \frac{1}{2^{2M}} R_M^{SQ} + \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} R_{m,d} \tag{18}$$

where R_M^{SQ} corresponds to the bit allocation, in bits per pixel, of lowest resolution M subimage.

Assignment of the weights is based on the fact that the human eye is not equally sensitive to signals at all spatial frequencies. On the basis of contrast sensitivity data collected by Campbell and Robson [10], and to obtain a controlled degree of noise shaping across the subimages, we consider a function $B_{m,d}$ such that:

$$B_{m,d} = \gamma^m \log(\sigma_{m,d}^{2\beta_{m,d}}) \tag{19}$$

where $\sigma_{m,d}$ is the standard deviation corresponding to subimage (m, d) and the values of γ and $\beta_{m,d}$ are chosen experimentally in order to match human vision.

$D_T^*(R_T)$ is the total weighted encoding distortion function, and M is the lowest resolution considered.

The expression of $D_{m,d}(R_{m,d})$ is given by [19]

$$D_{m,d}(R_{m,d}) = 2^{-cR_{m,d}} \times \alpha_{m,d}(p, c), \qquad c \geq 1$$

with

$$\alpha_{m,d}(p, c) = A(k_{m,d}, c) \times \left[\int [p_{m,d}(x)]^{k_{m,d}/(c + k_{m,d})} dx \right]^{(c + k_{m,d})} \tag{20}$$

This minimization problem can be solved by using Lagrangian multipliers. Using this technique, we must solve the following equation:

$$\frac{\partial}{\partial R_{m,d}}\left[D_T^*(R_T) - \lambda\left(R_T - \frac{1}{2^{2M}} R_M^{SQ} - \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} R_{m,d}\right)\right] = 0 \tag{21}$$

where λ is a Lagrangian multiplier.

Using (17) and (20), this equation becomes:

$$\frac{\partial}{\partial R_{m,d}}\left[\frac{1}{2^{2m}} D_M^{SQ}(R_M^{SQ}) + \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} (2^{-cR_{m,d}} \alpha_{m,d}(p, c) B_{m,d}) - \lambda\left(R_T - \frac{1}{2^{2M}} R_M^{SQ} - \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} R_{m,d}\right)\right] = 0. \tag{22}$$

Taking the partial derivative with respect to $R_{m,d}$ yields an expression for $R_{m,d}$ in terms of λ:

$$R_{m,d} = \frac{1}{c} \log_2 \left[\frac{(c \ln 2)\alpha_{m,d}(p, c) B_{m,d}}{\lambda}\right]. \tag{23}$$

By substituting (23) into the constraint (18) of the minimization problem we obtain an expression of the Lagrangian multiplier λ

$$\lambda = c \ln 2 \left[2^{-c(R_T - (1/4^M) R_M^{SQ})} \prod_{m=1}^{M} \prod_{d=1}^{3} \cdot [\alpha_{m,d}(p, c) B_{m,d}]^{1/4^m}\right]^{4^M/4^M - 1}. \tag{24}$$

Finally, substituting λ into (23) results in an expression of the optimal bit assignment $R_{m,d_{opt}}$ (in bits per pixel (bpp)) to the vector quantizer of subimage (m, d):

$$R_{m,d_{opt}} = \frac{4^M R_T - R_M^{SQ}}{4^M - 1} + \frac{1}{c} \log_2 \left[\frac{\alpha_{m,d}(p, c) B_{m,d}}{\left[\prod_{m'=1}^{M} \prod_{d'=1}^{3} [\alpha_{m',d'}(p, c) B_{m',d'}]^{1/4^{m'}}\right]^{4^M/4^M - 1}}\right]. \tag{25}$$

This expression requires the knowledge of the subimage's PDF's.

The optimal distortion of the quantizer, $D^*_{T_{opt}}(R_T)$, is then computed by combining (25) and (17). We find:

$$D^*_{T_{opt}}(R_T) = \frac{1}{2^{2M}} D_M^{SQ}(R_M^{SQ}) + \frac{4^M - 1}{4^M} 2^{-c(4^M R_T - R_M^{SQ})/4^M - 1} \cdot \left[\prod_{m=1}^{M} \sum_{d=1}^{3} [\alpha_{m,d}(p, c) B_{m,d}]^{1/4^m}\right]^{4^M/4^M - 1}. \tag{26}$$

Finally, bit allocation which is a function of the image will be transmitted as side information requiring only a few bits.

IV. Experimental Results

The images used are sampled 256 by 256 black and white images. The intensity of each pixel is coded on 256 grey levels (8 bpp).

The numerical evaluation of the coder's performance is achieved by computing the peak signal-to-noise ratio (PSNR) between the original image and the coded image.

For each coded image, we can use a variable length code. We also give the corresponding $\mathcal{R}_T$ if an optimal entropy coding was performed, defined as follows.

To the L codewords w_j; $j = 1, 2, \cdots, L$ of the vector quantizer corresponds to L regions (clusters) of $\mathbb{R}^k$, $\mathcal{P}_j$; $j = 1, 2, \cdots, L$. The jth region is defined by

$$\mathcal{P}_j = \{x \in \mathbb{R}^k / Q(x) = w_j\}$$

and represents the subset of vectors of $\mathbb{R}^k$ which are well matched by the codeword w_j of the codebook.

Thus for each resolution and direction, we can introduce the average information of the codebook, called the entropy measure:

$$\mathcal{R}_{m,d} = -\frac{1}{k_{m,d}} \times \sum_{j=1}^{L} p(w_j) \log_2 p(w_j) \text{ bpp}$$

where $p(w_j)$ is the probability of selecting the source vector w_j, belonging to the codebook at scale m and corresponding to the orientation d, during the coding of the image (m, d).

Then, as in (18), $\mathcal{R}_T$ is the sum of the estimated entropy in each subimage as follows:

$$\mathcal{R}_T = \frac{1}{2^{2M}} \mathcal{R}_M^{SQ} + \sum_{m=1}^{M} \frac{1}{2^{2m}} \sum_{d=1}^{3} \mathcal{R}_{m,d} \text{ bpp}.$$

The vector quantizer used is a *full search* quantizer, i.e., during the coding, all of the vectors in the subcodebook corresponding to the resolution and direction to be encoded are searched. The selection criterion used is the MSE criterion.

A. Comparison Between the Different Wavelets

In the following, we present results obtained with the Lena image (image within the training set) for a real bit rate of 1 bpp and using the three different filters proposed in Section II-B. (Fig. 13 corresponds to filters 9-3 presented in example 1, Fig. 14 corresponds to filters 9-7 presented in example 2, and Fig. 15 corresponds to filters 5-7 presented in example 3.) Here, the Lena image is taken as part of the training set in order to minimize the effects of quantization noise: this enables the influence of the filters to be taken into account.

For a given set of filters, separate codebooks are trained for each resolution-orientation subimage, and bit alloca-

Fig. 13. Filters no. 1 9-3, PSNR = 31.82 dB, $\mathcal{R}_T = 0.80$ bpp.

Fig. 15. Filters no. 3, 5-7, PSNR = 31.46 dB, $\mathcal{R}_T = 0.80$ bpp.

Fig. 14. Filters no. 2, 9-7, PSNR = 32.10 dB, $\mathcal{R}_T = 0.78$ bpp.

Fig. 16. Original 256 by 256 Lena, 8 bpp.

tion is carried out according to (25). For the Lena image, the bit assignment is represented in Fig. 17. Resolution 1 (diagonal orientation) is discarded. Resolution 1 (horizontal and vertical orientations) and resolution 2 (diagonal orientation) are coded using 256-vector codebooks (codeword size 4 by 4) resulting in a 0.5-b/pixel rate, while resolution 2 (horizontal and vertical orientations) is coded at a 2-b/pixel rate using 256-vector codebooks (codeword size 2 by 2). Finally, the lowest resolution is coded at 8 b/pixel.

B. Results as a Function of Regularity and Vanishing Moments

In Section II-B, we mentioned our belief that both the regularity of the reconstruction wavelet $\tilde{\psi}$ and the number

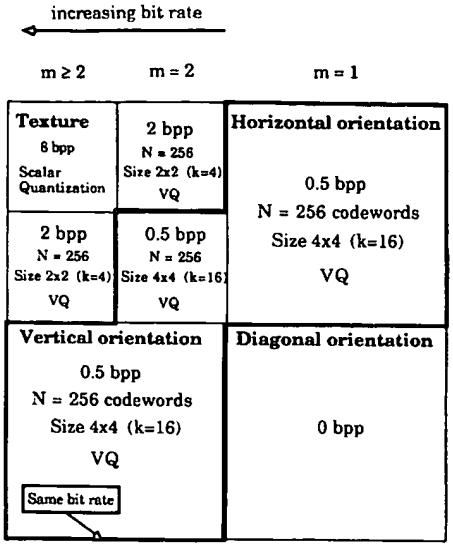


Fig. 17. Subimages bit rate allocation: example of a bit allocation for a total bit rate of 1 bpp and for the 256 by 256 Lena image.

of vanishing moments of the analyzing wavelet ψ are important in applications. To illustrate this we carried out the following experiments. For a given pair, h, $\tilde{h}$, we analyzed the same image twice: once as described above, and a second time after exchanging the roles of the filters h and $\tilde{h}$.

The filter pairs in example 2 both have the same number of vanishing moments, $k = \tilde{k} = 4$. However, $\tilde{\psi}$ is considerably more regular than ψ (see Fig. 3). With this filter pair, our experiment on the Lena image led to a PSNR of 32.10 dB in the first case, and to a PSNR of 31.51 dB if the roles of h and $\tilde{h}$ are inverted. The case where the reconstruction wavelet has the highest regularity, therefore, performs best.

In example 1 the functions ψ and $\tilde{\psi}$ have comparable regularity: both are continuous and neither has a continuous derivative. In fact $\tilde{\psi}$ is a bit more regular than ψ: $\tilde{\psi}$ is differentiable almost everywhere, and is Hölder continuous with exponent 1, while ψ is Hölder continuous with the exponent only at 0.83. On the other hand, ψ has 2 vanishing moments, while $\tilde{\psi}$ has 4 ($k = 4$, $\tilde{k} = 2$). The same experiment, again with the Lena image, now leads to a PSNR of 31.82 dB if h, $\tilde{h}$ are taken as in Table I, and to a PSNR of 31.13 dB when the roles of h and $\tilde{h}$ are reversed. The situation where $\tilde{\psi}$ is most regular but ψ has fewer vanishing moments, therefore, performs better (gain of 0.69 dB) than the case where ψ has more vanishing moments but $\tilde{\psi}$ is less regular. This seems to suggest that the regularity of $\tilde{\psi}$ has a larger effect than the number of vanishing moments of ψ. However, in this example the difference in overall regularity, as measured by the differences between Hölder exponents, is much smaller here than in example 2 (0.17 as compared to 0.63 in example 2), and it seems hard to explain how this smaller difference in Hölder exponent could account for a comparable gain in PSNR. In fact, the Hölder exponent is not a very good measure for the regularity of $\tilde{\psi}$ in this case: it is completely determined by the discontinuity of the derivative of $\tilde{\psi}$ in only a few points, and it is insensitive to the fact that $\tilde{\psi}$ is infinitely differentiable in all other points. If this is taken into account, then $\tilde{\psi}$ looks much more regular than ψ (the Hölder exponent of which is determined by its behavior near a dense set of points), which might explain the gain in PSNR.

We conclude from all this that: 1) for the same number of vanishing moments for ψ, the scheme with most regular $\tilde{\psi}$ is likely to perform best; and 2) increasing the regularity of $\tilde{\psi}$, even at the expense of the number of vanishing moments for ψ, may lead to better results.

Based on theoretical arguments (Taylor expansions) and results from numerical analysis [8], we also expect: 3) for comparable regularity of $\tilde{\psi}$, the scheme with largest vanishing moments for ψ is likely to perform best.

C. Comparison with Other Coders

If the PSNR is chosen as a criterion of comparison, these results are close to those obtained by Woods and O'Neil [42] and Westerink *et al.* [40]. However, in their subband coding algorithm, they use 32-taps Johnston filters, while only 9 or 7 taps are necessary for our method. According to Westerink's results in [41], the PSNR decreases by about 2 dB when using 8-taps Johnston filters. However, some others new QMF designs can also lead to good results with about 9 taps for image coding [1].

In this section, we present both numerical and qualitative comparison between our coding scheme and other previously published results. Since the most popular image in the recent literature has been the 512 by 512 Lena image, the comparison is made using this image taken outside the training set.

Among the different methods published, we consider the three following well-known methods: Ho and Gersho obtained a 30.93-dB PSNR at 0.36 bpp, result using "variable-rate multi stage VQ" [23]. Riskin and Gray improved on the full search VQ (PSNR = 29.29 dB, 0.32 bpp) using pruned tree structured VQ (PSNR = 30.92 dB, 0.32 bpp) [34]. High PSNR values were obtained by Woods and Cohen using entropy coded and predictive VQ (PSNR = 32.5 dB, 0.45 bpp) [11].

Our aim is not to optimize the PSNR but rather a weighted function of the MSE in order to match human vision. We give two examples at low bit rate using wavelet VQ.

Our initial result at 0.37 bpp presented Fig. 18 with a 30.85-dB PSNR is very close to those of Ho and Gersho [23] and Riskin *et al.* [34]. The perceptual quality of our coded images is better than indicated by the PSNR value

Fig. 18. 512 by 512 Lena image. Filters no. 2 9-7, PSNR = 30.85 dB. $\mathcal{R}_T = 0.37$ bpp.

Fig. 19. 512 by 512 Lena image. Filters no. 2 9-7, PSNR = 29.11 dB. $\mathcal{R}_T = 0.21$ bpp.

mainly due to the regularity of the wavelet and the bit allocation. These images do not suffer from the blocking effects obtained when using VQ in the spatial domain. No ringing effects can be observed.

The second result at 0.21 bpp presented in Fig. 19 with a 29.11-dB PSNR shows that a very low bit rate can be achieved with our method, without severe degradation.

Our method using a new class of filters derived from wavelet theory using full search VQ can be improved by any of the three above-mentioned methods.

In fact the LBG clustering algorithm is a very simple algorithm but not optimal for variable length code. The PSNR of the method could be improved by about 3 dB, for example, using ECVQ [34] but CPU time becomes prohibitively expensive.

D. Progressive Transmission Scheme

The main objective of progressive transmission is to allow the receiver to recognize a picture as quickly as possible at minimum cost, by sending a low resolution level picture first. Then, it can be decided to receiver further picture details or to abort the transmission. Further details of the picture are obtained by sequentially receiving the encoded wavelet coefficients at different resolution levels and directions.

Following the example of [40], we will display each picture level during the progressive transmission with a size that matches the resolution of that particular level.

To test the efficiency of the vector quantizer, the image to be coded is taken outside the training set.

Fig. 20 represents 5 stages in the progressive transmission of a 256 by 256 image using filters 9-7 given in example 2. According to the bit allocation procedure (Section III-C) with a generalized Gaussian PDF approximation law, only the wavelet coefficients corresponding to the $m = 1$ and $m = 2$ high resolution levels are vector quantized, while the low level subimages ($m \geq 2$) are scalar quantized.

V. Conclusion

This paper describes a new image coding scheme combining the wavelet transform and VQ.

A new family of filters has been derived from the wavelet theory. We have shown the importance of regularity and vanishing moments for image coding. Furthermore, these filters require few taps, unlike standard QMF methods.

The wavelet transform used here attempts to exploit the masking effect of the human eye, yielding encouraging results. Indeed, the proposed method enables high compression bit rates while maintaining good visual quality through the use of bit allocation in the subimages. The blocking effects seen when spatial VQ is performed are avoided.

This method is well adapted to progressive transmission as well as very low bit rate compression. Furthermore, using a simple full-search VQ provides good results, comparable to the best results published currently.

Further research should include some new derivation such as entropy constraint and predictive VQ. We would improve this coding scheme, if we accept a heavier computational load.

Resolution	*m=4*	*m=3*	*m=2*	*m=1*
Size	16 × 16 pix	32 × 32 pix	64 × 64 pix	128 × 128 pix
R_T	0.031 bpp	0.125 bpp	0.5 bpp	0.781 bpp
$\mathcal{R}_T$	0.0264 bpp	0.0919 bpp	0.3354 bpp	0.5039 bpp

Resolution m=0

256 × 256 pix

$R_T = 1$ bpp $\quad \mathcal{R}_T = 0.6297$ bpp

PSNR = 31.28 dB

Fig. 20. Progressive transmission-filters no. 2 9-7.

REFERENCES

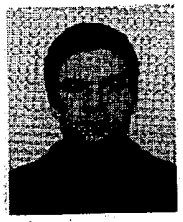
[1] E. H. Adelson and E. Simoncelli, "Non-separable extensions of quadrature mirror filters to multiple dimensions," *Proc. IEEE*, vol. 78, Apr. 1990.

[2] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.

[3] V. R. Algazi, "Useful approximation to optimum quantization," *IEEE Trans. Commun.*, vol. COM-14, pp. 297-301, June 1966.

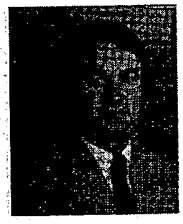
[4] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image

coding using vector quantization in the wavelet transform domain," in *Proc. IEEE ICASSP*, April 1990, pp. 2297-2300.
[5] M. Barlaud, L. Blanc-Féraud, P. Mathieu, J. Menez, and M. Antonini, "2D linear predictive image coding with vector quantization," in *Proc. EUSIPCO*, Grenoble, France, Sept. 5-8, 1988, pp. 1637-1640.
[6] M. Barlaud, P. Mathieu, and M. Antonini, "Wavelet transform image coding using vector quantization," presented at 6th Workshop on MDSP, Monterey, CA, Sept. 1989.
[7] G. Battle, "A block spin construction of wavelets. Part I Lemarié functions," *Comm. Math. Phys.*, vol. 110, pp. 601-615, 1987.
[8] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical analysis. I," to be published.
[9] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, pp. 482-540, 1983.
[10] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Phys.*, vol. 197, pp. 551-566, 1968.
[11] R. A. Cohen and J. W. Woods, "Sliding block entropy coding of images," in *Proc. IEEE ICASSP*, Glasgow, Scotland, May 23-26, 1989, pp. 1731-1733.
[12] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," AT&T Bell Lab., Tech. Rep., TM 11217-900529-07, 1990.
[13] J. H. Conway and N. J. A. Sloane, "A lower bound on the average error of vector quantizers," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 106-109, Jan. 1985.
[14] I. Daubechies, A. Grossman, and Y. Meyer, "Painless nonorthogonal expansions," *J. Math. Phys.*, vol. 27, pp. 1271-1283, 1986.
[15] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," to be published.
[16] —, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909-996, 1988.
[17] —, "Orthonormal bases of compactly supported wavelets. II. Variations on a theme," AT&T Bell Lab., Tech. Rep. TM 11217-891116-17, 1990.
[18] J. C. Feauveau, "Analyse multirésolution par ondelettes non orthogonales et bancs de filtres numériques," Ph.D. dissertation, Univ. Paris Sud, France, Jan. 1990.
[19] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, July 1979.
[20] —, "On the structure of vector quantizers," *IEEE Trans. Inform. Theory*, vol. IT-28, Mar. 1982.
[21] R. M. Gray, "Vector quantization," *IEEE ASSP Mag.*, pp. 4-29, Apr. 1984.
[22] A. Grossman and J. Morlet, "Decomposition of hardy functions into square integrable wavelets of constant shape," *SIAM J. Math Anal.*, vol. 15, pp. 723-736, 1984.
[23] Y. Ho and A. Gersho, "Variable-rate multi-stage vector quantization for image coding," in *Proc. IEEE ICASSP*, New York, Apr. 1988.
[24] P. G. Lemarié, "Une nouvelle base d'ondelettes de $L^2(\mathbb{R})$," *J. Math. Pures et Appl.*, vol. 67, pp. 227-238, 1988.
[25] P. G. Lemarié and Y. Meyer, "Ondelettes et bases hilbertiennes," *Rev. Mat. Iberoamericana*, vol. 2, pp. 1-18, 1986.
[26] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84-95, Jan. 1980.
[27] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 11, July 89.
[28] D. Marr, *Vision*. New York: Freeman, 1982.
[29] P. Mathieu, M. Barlaud, and M. Antonini, "Compression d'Images par transformée en ondelette," *12ième colloque GRETSI, Juan les Pins*, June 12-16, 1989.
[30] P. Mathieu, M. Barlaud, and M. Antonini, "Compression d'Image par transformée en ondelette et quantification vectorielle," *Traitement du Signal*, vol. 7, no. 2, 1990.
[31] Y. Meyer, "Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs," *Seminaire Bourbaki*, no. 662, 1985-1986.
[32] N. M. Nasrabadi and R. A. King, "Image coding using vector quantization: A review," *IEEE Trans. Commun.*, vol. 36, Aug. 1988.
[33] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978.
[34] E. Riskin, E. M. Daly, and R. M. Gray, "Pruned tree-structured vector quantization in image coding," in *Proc. IEEE ICASSP*, Glasgow, Scotland, May 1989, pp. 1735-1738.
[35] M. J. Smith and D. P. Barnwell, "Exact reconstruction for tree-structured subband coders," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-34, pp. 434-441, 1986.
[36] J. O. Stromberg, "A modified haar system and higher order spline systems," in *Conf. in Harmonic Analysis in Honor of Antoni Zygmund*, Vol. II, pp. 475-493.
[37] M. Vetterli, "Splitting a signal into subsampled channels allowing perfect reconstruction," in *Proc. IASTED Conf. Appl. Signal Processing Digital Filtering*, Paris, France, June 1985.
[38] M. Vetterli and C. Herley, "Wavelets and filter banks: Relationships and new results," in *Proc. IEEE ICASSP*, Albuquerque, Apr. 1990.
[39] P. H. Westerink, D. E. Boekee, J. Biemond, and J. W. Woods, "Subband coding of image using vector quantization," *IEEE Trans. Commun.*, vo. 36, pp. 713-719, 1988.
[40] P. H. Westerink, J. Biemond, and D. E. Boekee, "Progressive transmission of images using subband coding," in *Proc. IEEE ICASSP*, 1989, pp. 1811-1814.
[41] P. H. Westerink, "Subband coding of images," Ph.D. dissertation Delft Univ., 1989.
[42] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-34, Oct. 1986.
[43] P. Zador, "Asymptotic quantization error of continuous signals and their quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 139-149, 1982.



Marc Antonini was born in France on August 29, 1965. He received the DEA degree in signal processing in 1988 from the University of Nice-Sophia Antipolis, France, and the Ph.D. degree from the Laboratory of Signaux et Systèmes, URA 13S, CNRS and the University of Nice-Sophia Antipolis in 1991.

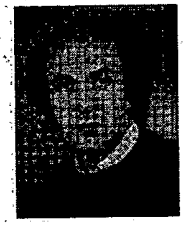
His research interests include multidimensional image processing, wavelet analysis, and image coding.



Michel Barlaud (M'88) was born in France on November 24, 1945. He received the "Doctorat d'Etat" degree from University of Paris XII.

He is currently a Professor and a member of the Laboratory of Signaux et Systèmes, URA 13S both from CNRS and University of Nice-Sophia Antipolis. After some work on non-stationary signal processing, his research interests move towards multidimensional image processing, wavelet analysis, image coding, inverse problems, image restoration, and edge detection.

Dr. Barlaud is member of the IEEE-ASSP MDSP committee.



Pierre Mathieu was born in Alger on May 10, 1956. He received the Ingenieur ENSEEIHT and Ph.D. degrees from INP Toulouse.

He is currently Maître de Conférences in the Laboratory of Signaux et Systèmes, URA 13S both from CNRS and University of Nice-Sophia Antipolis. His research interests include multidimensional image processing, wavelet analysis, image coding, and image restoration.

Ingrid Daubechies (M'89) received the B.S. and Ph.D. degrees from the Vrje Universiteit Brussel, Belgium in 1975 and 1980, both in physics.

She is currently a Member of Technical Staff in the Mathematics Center of AT&T Bell Laboratories, Murray Hill, NJ. Her current research interests include mathematical problems in connection with signal analysis, in particular applications of time-frequency representations.